②

AD

AD-E402 350

Technical Report ARPAD-TR-92002

SPREADSHEET-LIKE IMAGE ANALYSIS

Paul Willson

August 1992

US ARMY
ARMAMENT MUNITIONS
& CHEMICAL COMMAND
ARMAMENT RDE CENTER

U.S. ARMY ARMAMENT RESEARCH, DEVELOPMENT AND ENGINEERING CENTER

Product Assurance & Test Directorate

Picatinny Arsenal, New Jersey

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE August 1992	3. REPORT TYPE AND DATES COVERED

4. TITLE AND SUBTITLE SPREADSHEET-LIKE IMAGE ANALYSIS	5. FUNDING NUMBERS
6. AUTHOR(S) Paul Willson	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESSES(S) ARDEC, QA Technology Office (SMCAR-QAH-T) Picatinny Arsenal, NJ 07806-5000	8. PERFORMING ORGANIZATION REPORT NUMBER Technical Report ARPAD-TR-92002
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(S) ARDEC, IMD STINFO BR (SMCAR-IMI-I) Picatinny Arsenal, NJ 07806-5000	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES	
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited	12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This report describes the design of a new software system being built by the Army to support and augument automated nondestructive inspection (NDI) on-line equipment implemented by the Army for detection of defective manufactured items. The new system recalls and post-processes (off-line) the NDI data sets archived by the on-line equipment for the purpose of verifying the correctness of the inspection analysis paradigms, of developing better analysis paradigms, and to gather statistics on the defects of the items inspected.

The design of the system is similar to that of a spreadsheet, i.e., an array of cells which may be programmed to contain functions with arguments being data from other cells and whose resultant is the output of that cell's function. Unlike a spreadsheet, the arguments and the resultants of a cell may be a matrix such as a two-dimensional matrix of picture elements (pixels). Functions include matrix mathematics, neural networks and image processing as well as those ordinarily found in spreadsheets. The system employs all of the common environmental supports of the Macintosh computer, which is the hardware platform. The system allows the resultant of a cell to be displayed in any of multiple formats such as a matrix of numbers, text, an image, or a chart. Each cell is a window onto the resultant. Like a spreadsheet if the input value of any cell is changed its effect is cascaded into the resultants of all cells whose functions use that value directly or indirectly. The system encourages the user to play what-if games, as ordinary spreadsheets do.

14. SUBJECT TERMS Image analysis, nondestructive inspection, spreadsheet, Macintosh software, neural network, signal processing, signal analysis, spectral analysis, matrix mathematics	15. NUMBER OF PAGES 14
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAR

NSN 7540-01 280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

CONTENTS

INTRODUCTION

The work described by this paper is relevant to automated inspection of manufactured products or components of products by nondestructive evaluation (NDE) systems. The manufacturing scenario of concern is that where a manufacturer produces a multitude of the same item; the item is made to some specification (target value and limited deviation from target); and a stable NDE system is used for inspecting the item. The evaluation process can be described as the analysis of relative differences in the NDE signal for the purpose of detecting and measuring by implication differences between items within the volume manufactured.

The Product Assurance Directorate of the U.S. Army AMCCOM has taken a leading role in the development of fully automated NDE systems. A significant difference between our NDE systems and more traditional ones is the archival of the raw signal. Until recently the cost of archiving the raw signal was prohibitive. Today, an EXABYTE tape drive (ref 1) can store up to five gigabytes of data onto a single 8mm tape in digital format at an unprecedented low cost of about 0.2¢/Mbyte. Under normal production many gigabytes of data may be archived in a single day. The archived signal can be used for improving NDE signal analysis.

NDE is based on the "seeing" of product characteristics through eyes which see visible light, infrared, microwave, x-ray, gamma ray, sound, ultrasound, electronic and magnetic properties, etc. Any of these NDE signals can be displayed as an image. The types of characteristics looked for can be classified by the signal strength and the spatial size and location of the characteristic, i.e. position and condition of components in assemblies, material internal structure such as voids, porosity, fissures, granularity, and density, surface characteristics such as scratches, dents, and topology.

The manual review of filmed x-ray images requires only a simple, inexpensive, and portable optical lightbox. Digitized x-ray images, as coming from automated NDE systems, also need to be manually reviewed, but obviously require more sophisticated equipment than a lightbox. As the number of automated NDE systems increases so will the need for digital image review systems. NDE systems may generate 1-D, 2-D, 3-D, or multi-dimensional image data. Examples of 1-D image data are signal amplitude versus time. Examples of 2-D are x-ray and ultrasound images. Examples of 3-D image data are x-ray tomography. Multi-dimensional image data would originate from the fusion of multiple NDE signals. Features and options are needed to make up for the lack of density and spatial resolution found in the digitized image and the dynamic range of the display medium (usually a CRT) as compared to that of film. These factors add to the complexity and ultimately the cost of digital image review systems.

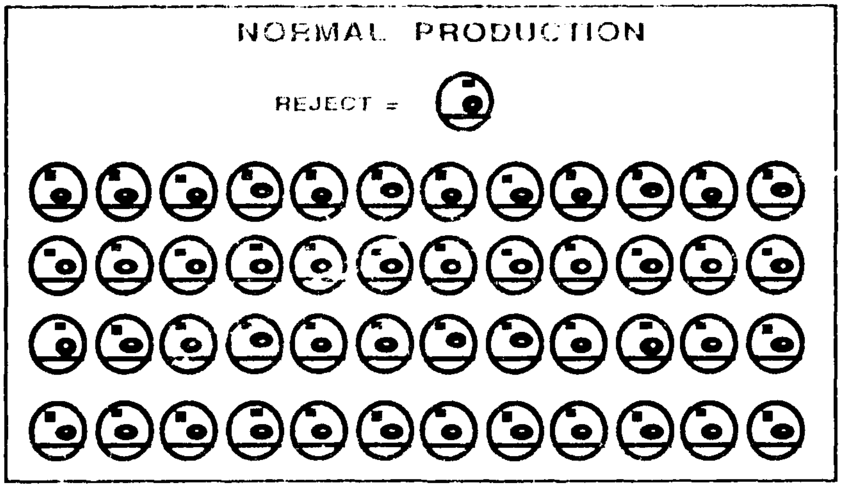

Figure Manual review of x-ray images involves "seeing" the characteristics indicative of a unacceptable product while ignoring the natural signal variations which are acceptable. It's left as an exercise to the reader to find the four rejects.

The accumulation of gigabytes of raw NDE signal opens new vistas to the NDE community. Four obvious uses for the accumulated data are: (1) verification of the proper functioning of the automated NDE equipment (quality of signal), (2) verification that the NDE algorithms are functioning properly (quality of evaluation), (3) independent evaluation of the quality of the manufactured product, and ultimately the (4) development of new or modified evaluation algorithms. In the past, NDE algorithm development most often had to be done with very limited and/or artificial data.

The verification processes and the development of automated NDE signal analysis algorithms entails a large amount of manual analysis. Manual analysis is basically a process of making characteristics of interest or importance in the product or in the signal visible to the user. Once the user "sees" the characteristics, he can create and judge the automated analysis (fig.). Manual analysis is used to categorize and reduce the large volume into subsets representing those characteristics of interest. Since the volume of data can be overwhelming and since the data is in digital format, the reduction process is best done by good generic semi-automated computer based tools, tools which reduce the labor and tedium.

Many scientists and engineers, representing many companies, universities and government, are working in the NDE field developing automated analysis algorithms including the application of artificial neural networks. It's been our experience that more time and energy is spent building and maintaining the development environment than is spent developing algorithms and analyzing the data. The tools required by each of these groups, whether simply reviewing image data or developing automated analysis algorithms, are of very similar nature. This project was

begun to build a generic, inexpensive digital image review and post-processing system applicable to NDE data of any dimension, a system which would not require the scientist to have special training nor special programming knowledge, a system which any group working this field would find useful. It's hoped the system will provide the "leap ahead" to NDE analysis that spreadsheets provided to the numeric analysis field. The following section will describe the underlying nature of review and analysis upon which a Generic Image Post-processing and Review (GIPR) system is being built.

NDE SIGNAL REVIEW AND ANALYSIS

A scientist or engineer manually reviewing and analyzing a NDE signal, does so by mentally relating the NDE signal or a "portion" of that signal to his concept of "quality." A mental relationship is usually made based on the degree of coincidence in time and space of two or more observed events. One of those events is likely to be an action performed or triggered by the person. Events in our case will include pictorial outputs from a computer. Performing this relationship might be termed man/machine synergism. A measure of the success of the GIPR system is the degree to which it encourages the user to try many alternatives until he is satisfied with the results, similar to the "what if" scenarios when using spreadsheets.

The GIPR system needs two distinct levels of operation: (1) those necessary to make data meaningful to the observer and (2) calculations which result in numeric values that need be saved or used by another calculation. The results of the first are temporal, disappearing when the CRT contents are changed. Due to the inherent spatial and density resolution limits of the CRT, making data meaningful to the observer often involves distorting the data nonlinearly while simultaneously displaying the data in its normal context. An example would be using a magnifying glass to see details and then looking around the magnifying glass to see the detail in the bigger picture.

An IBM study (ref 2) in the 1970's by Walter J. Doherty and Richard P. Kelisky found a one-to-one increment in user think time as system response time improved. Later studies at IBM, when 300 ms system response times were possible, found even more dramatic increases in productivity as system response time improved. The results of one of these studies (ref 3), shows the productivity gains up to a factor of nine for an expert and up to three for an average group as response time decreased to 300 ms. The processes under study by IBM and the processes GIPR will be used for are parallel in nature. One of GIPR's intent is to decrease the time between an idea and its implementation dramatically over what is generally possible in order to increase the user's productivity.

A study by MacLachlan, Newsome and Shattuck (ref 4) has shown that people can read and comprehend information 15 to 20% faster when using a high

resolution CRT than when using a low resolution CRT. Since the GIPR system is being explicitly designed so the scientist can "see" the impact of his actions, the system should employ the highest resolution display possible. In addition, since there will be numerous diverse relationships which the user must see, the GIPR system will need a large display area and optimally use that area. Use of multiscreens and windows is the best means of achieving large area and optimal use thereof.

Development of NDE algorithms proceeds more or less as follows: (1) manual assessment of examples - real or artificial, (2) composition of algorithm, (3) processing of examples by the algorithm, (4) review of results, and (5) iteration of steps 1 through 4 until convergence with satisfactory results is attained. We have found that thousands of examples are required for developing good NDE algorithms whether they be conventional analysis algorithms or artificial neural networks. The process is labor intensive and costly. The accuracy of the manual assessment and the accuracy and precision of the mathematical calculations in the algorithm is paramount to the success of the process.

The ability to "see" the results will depend on having an appropriate display format for the results. Since the magnitude and type of relationships may be very diverse, the GIPR system needs to provide a large and diverse selection of display formats. It should provide a meaningful and quick way of displaying intermediate results. Some programs which display data pictorially lack precision in their rendition. In this case all results and values must be displayed accurately and precisely. The system needs to give the scientist the ability to see the details and simultaneously see the "big" picture.

The system should provide a simple and quick method of composing and editing algorithms. The system should provide an environment for developing algorithms based on selected examples chosen from the many and then extending the process to the many.

The acquisition of NDE signal is basically a serial process which follows a serial manufacturing process. Granted, parallel operations may be going on, but at least on a large scale the overall operation can be viewed as serial. This is the basis for the concept for "lot acceptance" by inspection of representative product and also is the basis for statistical process control. For a well behaved manufacturing process one should expect acceptable natural variation from product to product and also trends over time in product characteristics. The inspection process must determine the anomalies and also see the trends. The trends give information necessary for keeping the process "in control," which will prevent the manufacturing of rejects. Keeping all of this mind, the GIPR system needs to view the NDE signal both as a series of short signals each representing one item and also as a continuous signal representing a continuously varying process. This implies that the GIPR system must

be able to handle massive data sets looking for trends and simultaneously access at random smaller nearly identical subsets of the massive data set.

The raw NDE signal contains information indicative not only of unacceptable quality of the item under inspection and unacceptable quality of the inspection system (which we will call signal), but also of natural and acceptable variation within the item and within the NDE system (which we will call noise). Appropriate algorithms for NDE analysis can be anything from the most simple to the most complex. A desired NDE algorithm will separate the unacceptable from the acceptable, or enhance the signal/noise ratio, or be sensitive to the signal and insensitive to the noise. It should also separate the signal into components indicative of the quality of the item from those indicative of the quality of the inspection system.

Ultimately the quality of the item is dependent on the quality of the manufacturing process which made it. In as much as a relationship between the item's quality, as expressed in variations of the NDE signal, can be established with particular variations in the manufacturing process, the results of the analysis can be used as feedback for process control. Process control requires the accumulation of statistics on the results of algorithms over the entire set of examples.

Our experience is that development of worthwhile NDE algorithms for real manufactured items requires thousands of real examples, both for development and for proveout of the algorithms. The GIPR system must be able to both sequentially and randomly access and to process the many examples rapidly. For a given item all examples should be analyzed by the same algorithm. The algorithm may consist of many complex serial operations which could be applied in parallel to the signal. Development of an algorithm proceeds most rapidly when the user has the option of selecting examples randomly and redundantly as desired. Proveout requires selection of the examples in sequence while accumulating statistics on the results.

SYSTEM CONCEPT

Many key concepts desired by users of the GIPR system are met by existing computerized spreadsheets. For example, a spreadsheet can be used without programming knowledge. The user never needs to enter data more than once, although the data may be used many times by many formulas. Formulas can vary from simple to complex. Any data or the resultant of any formula can be fed forward as the argument to any other formula. The spreadsheet automatically recalculates any formula that depends on a value that has been changed. Each cell can be individually formatted for optimum meaning to the user. The feedback mechanism to the user following data and formula changes is natural and immediate. Playing "what if" games or experimenting with different formulas is natural.

There are some significant differences between the type of data input, the display formats, and the formulas required for NDE data analysis and those commonly found in spreadsheets. Instead of dealing with single values or elements of an array as in spreadsheets, the GIPR system must treat the array, or an array subset, as input data, as arguments for formulas, and/or as display elements. Whereas the input data in a spreadsheet is usually typed in or copied into cells, one value per cell, the input data to GIPR is expect to be that archived by automated inspections systems, i.e. multiple large arrays of 8 bit, 16 bit , or 24 bit signed or unsigned integers. Each array will be the data representing the NDE signal from one item. Since each array may be thousands of bytes, the user will not want to type or even copy this data into cells. The user will want to develop a formulation using the data of one item as an example and then check its correctness on many items. Output display formats needs to include those relevant to the data, i.e., pictures and arrays will replace the spreadsheet cell output. Statistical results over entire sets of items need to be accumulated and displayed.

The GIPR system is being designed to capitalize on those spreadsheet concepts which are useful while simultaneously accommodating the unique character of the NDE input data. The following similarities will exist between the GIPR system and spreadsheets.

1. A spreadsheet is composed of a matrix of cells with interrelationships defined by a formula within each cell whose resultant is displayed in the cell. A document in the GIPR system will be composed of multiple windows with interrelationships defined by a formula within each window whose resultant will be displayed in that window.

2. The user will never need to enter data more than once. Once entered, data can be used as an argument in any appropriate function in any window.

3. Formulas in a GIPR document can be complex formulation of many functions with many arguments.

4. The resultant matrix in a window of the GIPR document (or a segment of the window called a region-of-interest or ROI) can be fed forward as input argument to a function in another window's formula as opposed to a single value in a cell of a spreadsheet.

5. A function's arguments can be designated by name or by window (or ROI within a window) similar to the method found in a spreadsheet.

6. The GIPR system will automatically recalculate any formula and change its output display whose argument has been changed directly or indirectly at the time the argument was changed.

7. Each window of the GIPR document can be individually formatted for optimum meaning, but the formatting, as described below, will be much more extensive than in a spreadsheet.

8. A spreadsheet "save" command saves the entire document: all data, formats and formulas. A GIPR "save" command will save the entire GIPR document: all data, all formulas, all windows, window relative positions, window formats, and window ROIs.

9. A spreadsheet "open" command starts up the saved document as it was when saved. A GIPR "open" command will start up the saved GIPR document

10. The GIPR system will use pull-down menus, default values, and other normal spreadsheet environment techniques.

The GIPR system will have the following characteristics which differ from what exists in common spreadsheets.

1. A formula's resultant in a window can be a single value as found in a spreadsheet or can be, a 1-D, 2-D, or N-Dimensional matrix of values as per the formula.

2. A ROI may be designated by the user through an encircling procedure such as pulling out ("rubberbanding") a rectangle, ellipse, or a random shape.

3. The functions for formulas will be selected to be those appropriate for image processing, matrix arithmetic, as well as functions normally found in a spreadsheet.

4. The GIPR system is intended to be used for viewing and analysis of inspection data coming from many similar items, i.e. processing the archived NDE signal. The archived signal is not expected to be in the internal format used by GIPR. GIPR will have an IMPORT command to read in the archived file and reformat it. Each file of archived data may be the data representing one or more inspected items. If the number is greater than one, than the file needs to be artificially divided into subdata sets each representing the results of inspection of one item. The definition of the subdata set (SDS) will be alterable via the IMPORT menu. A special SDS window will be created by the IMPORT command to display the SDSs as individual items. The SDS window will differ in appearance from the others in that it will contain a scroll bar, instead of a function line, for scrolling through the SDSs or randomly selecting a SDS of the imported file. The SAVE command will save the SDS window along with all the others.

5. An extensive list of possible window formats including the SDS window will be available. A window's format can be changed to any "reasonable" format at any time. The resultant of a window's formula or the SDS may be displayed as:

a. A chart, such as a line, bar, dot, area, pie chart, scatter diagram, etc.

b. A 2-D intensity plot, i.e. a picture,

c. An overlay of a transparent picture on top of another picture,

d. A 3-D plot, line drawing or intensity plot w/hidden point removal and perspective control,

e. Numbers, displaying the resultant values in a matrix layout in the same relative locations as the image pixels would have, and

f. Text, displaying the ASCII interpretation of the resultant values.

Window formatting and control will not affect the resultant values; only the portrayal of the resultant values. The resultant whether it be a picture, chart, or text may extend outside the window borders. Scroll bars attached to the windows will allow one to scroll any portion of the resultant into the displayed space. Each window will have its own selectable magnification factor. Magnification of text will be done by selecting the font. Magnification of pictures and graphs will be done either by pixel replication in factors of two or by continuous stretching with pixel interpolation. Aspect ratio may be chosen as a fraction or be continuously changed by stretching. The possible display color values will be designated by the user by altering the hardware Look-Up-Table (LUT). Each picture window will have its own unique alterable software LUT to control the mapping of resultant values into the hardware LUT. Automatic density stretching will be optional and implemented by the software LUT. Color selection for charts will depend on the type of data being displayed.

Formulas can be composed of simple mathematical functions such as add, subtract, multiply and divide or highly specialized functions. Common matrix functions will be implemented such as matrix add, multiply, convolve, invert, and filter. Image transform functions such as erode, dilate, log, elevation, histogram, rotate, and Fourier Transform will be implemented. Decision making functions (categorizing, selection) such as applying an artificial neural network, maximum, minimum, greater that, less than, and equal to will be implemented. Measurement functions such as center of density, center of elevation, distance, angle, and sum will be implemented. Statistical functions such as standard deviation, norm, average, mean, and variance will be implemented.

CURRENT STATUS

An X-windows demonstrable prototype has been developed. A full-up system based on experience with the prototype is under development. The prototype was used for analyzing x-ray SDSs of 5,000 Army ammunition fuzes. Each SDS was composed of two x-ray views of 512 by 480 bytes or pixels. Altogether 2.5 gigabytes of data were analyzed. The data had been archived by an automated inspection system which had automatically categorized the quality of the fuzes. The same fuzes had been manually inspected from x-rays made on paper. Even so, interpretation of the archived data was non-trivial because the purpose of the analysis was to determine the quality of both the manual and the automated inspection and because manual analysis of the fuse images had never been done before by the individual doing it. As a result the individual had to learn the proper interpretation of information in the images.

Variation in image quality and product quality fluctuated almost randomly from fuze to fuze but in a statistical fashion over all. The image to image spatial and density variations arising from the acquisition system and the resolution of the system were on par with variations being looked for. For example a cubical shaped component in the fuze was found skewed in position relative to its surroundings. Excessive skewing became a discerning characteristic. The system's display parameters had to be "tuned" by trial and error over much of the data set in order for the user to "see" those characteristics which uniquely identified excessive skew of the component. Once the characteristics could be discerned, the entire data set had to be screened to tabulate statistics on the occurrence of the skewing. The user found it very important to be able to develop analysis in a trial and error fashion on one item and then to check the analysis on a random number of items and finally on the entire set. Classifying, categorizing and sorting was a large portion of the work. The prototype's ability to display the raw data, intermediate results and final results proved useful.

CONCLUSION

Future NDE systems should archive the raw NDE signal. The archived data will provide a tremendous aid in the development of good automated NDE algorithms. The described system is quite versatile, having the capability to analyze one image, or thousands of images with equal ease, play what-if games, etc. Scientist and engineers working in the NDE field should find the described system will reduce their labor and the tedium of NDE algorithm development.

A patent application is being made for the design concepts.

REFERENCES

1. EXABYTE Corporation, Boulder, CO.

2. Doherty, W.J. and Kelisky, R.P., IBM System J., Vol 18, No. 1, pp 143-163, 1979.

3. Brady, James T., A Theory of Productivity in the Creative Process, IEEE Computer Graphics and Applications, Vol 6, No. 5, pp 25-34, May 1986.

4. Better Computer Monitors Productivity, Rensselaer - The Alumni Magazine, Rensselaer Polytechnic Institute, Troy, NY, p 24, Dec 1989.

DISTRIBUTION LIST

Commander
Armament Research, Development and Engineering Center
U.S. Army Armament, Munitions and Chemical Command
ATTN: SMCAR-IMI-I (5)
SMCAR- QAH-T (5)
Picatinny Arsenal, NJ 07806-5000

Commander
U.S. Army Armament, Munitions and Chemical Command
ATTN: AMSMC-GCL (D)
Picatinny Arsenal, NJ 07806-5000

Administrator
Defense Technical Information Center
ATTN: Accessions Division (12)
Cameron Station
Alexandria, VA 22304-6145

Director
U.S. Army Material Systems Analysis Activity
ATTN: AMXSY-MP
Aberdeen Proving Ground, MD 21005-5066

Commander
Chemical Research, Development and Engineering Center
U.S. Army Armament, Munitions and Chemical Command
ATTN: SMCCR-MSI
Aberdeen Proving Ground, MD 21010-5423

Commander
Chemical Research, Development and Engineering Center
U.S. Army Armament, Munitions and Chemical Command
ATTN: SMCCR-RSP-A
Aberdeen Proving Ground, MD 21010-5423

Director
Ballistic Research Laboratory
ATTN: AMXBR-OD-ST
Aberdeen Proving Ground, MD 21005-5066

Chief
Benet Weapons Laboratory, CCAC
Armament Research, Development and Engineering Center
U.S. Army Armament, Munitions and Chemical Command
ATTN: SMCAR-CCB-TL
Watervliet, NY 12189-5000

Commander
U.S. Army Rock Island Arsenal
ATTN: SMCRI-TL/Technical Library
Rock Island, IL 61299-5000

Director
U.S. Army TRADOC Systems Analysis Activity
ATTN: ATAA-SL
White Sands Missile Range, NM 88002